DOCUMENT RESUME

ED 216 432 EA 014 604

AUTHOR Marshall, Catherine
TITLE Organizational and Political Tensions in Evaluation.
PUB DATE Mar 82
NOTE 48p.; Paper presented at the Annual Meeting of the American Educational Research Association (New York, NY, March 19-23, 1982).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Case Studies; Decision Making; Early Childhood Education; Educational Policy; *Evaluation Methods; Formative Evaluation; Legislators; Policy Formation; *Program Evaluation; *Research Reports; *Research Utilization; State Departments of Education
IDENTIFIERS California; *Interim Report on Study of Selected ECE Schools

ABSTRACT

After several lengthy evaluations of California's Early Childhood Education (ECE) program proved inadequate, unhappy legislators and educators finally found satisfaction in a brief, mimeographed report. Several factors were responsible for the success of this new evaluation, the "Interim Report on the Special Study of Selected ECE Schools with Increasing and Decreasing Reading Scores" (1977), published by the California State Department of Education. A pet project of the state superintendent, ECE had been subjected to nonthreatening evaluation processes only. Evaluation questions had never been clearly spelled out, partly because the interests of legislators and educators required different emphases. The successful evaluation occurred when the state department of education accepted the need for a study aimed at solving real problems, enabling evaluators to focus the interests of legislators and educators alike on one basic evaluation question while maintaining support from all involved parties. The evaluation's case study format permitted legislators to understand the research without extensive technical knowledge. The final report clearly explained the evaluation methodology, clarifying the generalizability and validity of the findings. The report's policy implications were direct and specific and, though aimed at all levels of the educational policy system, were framed so as to avoid organizational resistance. (Author/PGD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED216432

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Catherine Marshall

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ORGANIZATIONAL AND POLITICAL TENSIONS IN EVALUATION

Catherine Marshall

University of Pennsylvania
Graduate School of Education

Presented at the annual conference of the

American Educational Research Association

March 1982

New York

EA 014 604

Organizational and Political Tensions in Evaluation

Introduction

Do policymakers use research and evaluations for decisionmaking? What research methodologies are preferred, how important is the timing and method of reporting? How can researchers stay attuned to the shifting information needs of policymakers? Who should determine the relevant questions, appropriate methodologies, and the policy implications--the researchers or the policymakers? Must researchers educate policymakers about methodology, relevance, and limitations of research?

Researchers and evaluators have puzzled over these questions (see especially Mitchell, 1979; Weiss, 1973, 1980; Patton, 1978; and Dickey, 1980). This research explored these questions in a case study of evaluation utilization. It focused on The Interim Report, an evaluation of California's Early Childhood Education Program (ECE) which stood apart from other evaluations of ECE. Unlike other ECE evaluations, The Interim Report was used, by legislators, by the State Department of Education (SEA) and local school districts (LEAs) in their policymaking. Unlike previous evaluations, it used a case study approach. This paper emphasizes the management, negotiations with organization members and policymakers, and reporting of The Interim Report, assuming that evaluation utilization is much dependent upon these variables. This was a significant emphasis in an era and in a program which was fraught with political tensions and when policymakers were frustrated by evaluation controversies. (This paper does not detail the ways in which the evaluation was used, nor does it detail the case study methodology, but forthcoming papers will have these emphases.)

Perspectives

Literature on organizations, politics, and decisionmaking in education guided this study. It assumed that ECE decisionmaking was affected by organizational structures, by political tensions, by incrementalism, satisficing, and standard operating procedures (Lindblom, 1968; Allison, 1971; Weiss, 1977; March and Simon, 1958).

The research assumed that policymakers request information, research, and evaluation not only to make informed decisions, but also to further their political and organizational careers (Drew, 1979; House, 1973). Bureaucrats and policymakers maintain the appearance of logical decisionmaking with evaluation and research. Research and evaluation utilization is a valuable symbolic gesture for building political and organizational careers (Edelman, 1964). On the other hand, bureaucrats and policymakers who have career investments in certain programs may attack any unfavorable findings. They may seek control over the framing of questions, the methodology, the choice of researchers, the method and timing of reporting (Cohen, 1970; Sroufe, 1977).

This research assumed that we have much to learn about research utilization. We know that the timing, the format, and the mode of the research or evaluation report may be key factors in utilization (Weiss, 1972; Rosenthal, 1978; Mitchell, 1980). Patton (1978) has emphasized that bargaining and communication between policymakers and researchers is crucial to utilization. Alkin et al. (1979) has identified behaviors and structures which affect utilization at the school site. Weiss (1980) has found that policymakers' use of social science research depends upon its validity (which included conformity to users' expectations) and upon its utility (for examination of feasible alternatives and for challenging the status quo). Hill (1980) has emphasized the importance of policymakers' participation in the NIE Title I evaluation

management. Cohen (1977), Zeigler and Valentine (1979), Marshall (1981) and McLaughlin (1974) have illustrated that evaluation may be used more for political reasons than for informing decisionmaking. Thus, research on evaluation utilization must identify the crucial mix of political, organizational and procedural processes which is essential for utilization.

Finally, this research assumed that policymakers and bureaucrats are more familiar with research and evaluation which is framed, conducted, and reported within the experimental research paradigm. They expect exact findings, estimations of certainty, determinations of the bottom line, estimations of costs versus benefits of alternative policies. Some researchers and evaluators attempt this Herculean task; some promise it but do not deliver. A small, but growing cadre politely refuse but propose alternate methodologies and flexible approaches to question-framing. Case studies and ethnographic research and evaluation may be more appropriate for research and evaluation which will be of use to policymakers and bureaucrats (Wilson, 1979; Parlett and Hamilton, 1976).[1] It closely parallels the information-gathering strategies of most policymakers. Its focus and detailed description parallel the testimony of constituents and interest groups at hearings and in policymakers' offices.

This research was guided by a framework which emphasized the effects of organizational and political processes on education decisionmaking and it saw the importance of analyzing the practical, methodological, and political elements which affect evaluation/research utilization.

Methodology

This paper presents a focused analysis from a larger case study which examined the intermixing of policy actors' concerns with the use, and misuse

of evaluation in ECE from 1972 to 1980.

Exploratory field study methodology was chosen to discover processes of decisionmaking, communication and power among policy actors and variables in research and evaluation utilization.

Data collection consisted of participant observation, interviewing, and analysis of documents. The initial phase entailed constructing the history and identifying the policy actors in ECE. Following phases included the researcher interviewing forty policy actors in various state, county, and local education policy roles and in research institutions.

A pattern emerged showing that ECE evaluations were political tools. However, The Interim Report evaluation appeared as an anomoly, thus an intriguing focus for analysis. This paper is derived from an analysis of the policy system constraints and the organizational dynamics which affected evaluation utilization. The analysis identifies crucial strategies for organizational and political diplomacy to promote evaluation utilization.

Significance and Implications

Policymakers and researchers cannot decide on how to define and assess "good" policy and programs. Does it mean cost-effective, politically expedient, or educationally sound? Researchers despair over the misuse of their work, evaluators are frustrated when they must negotiate their design and play the diplomat with policymakers. Policymakers are tired of expensive evaluations which have uncertain findings, which come too late and fail to make linkages between research and policy.

This paper presents research on research. It details an instance of good evaluation utilization for policy and program planning. It identifies key processes which facilitate utilization. Evaluators and researchers who examine

these key processes can draw implications for evaluation management and for determinations of appropriate roles in the policymaking process for researchers. Policymakers can examine this case of evaluation utilization to identify the key political and organizational arrangements which may lead to more timely, relevant and useable policy information and to see how their actions constrain researchers.

Research findings such as these are important in an era when (1) legislatures are using evaluations in order to decide whether or not to "sunset" major educational and social programs, and state and local decisionmakers are setting priorities for funding through block grant procedures; (2) institutions of education research are losing support and credibility; (3) policymakers still seek practical, relevant and unbiased information to at least provide them with a perspective for viewing the information from vested interests; (4) states are taking a larger role in education policy decisions so we need case studies of their decision processes; and (5) evaluators are exploring alternative methodologies so we need to know how to convince policymakers of their worth.

The Findings: What was Special about this Evaluation?

The Interim Report evaluation differed from other evaluations of ECE in its question-framing, its methodology, its personnel, the political support, the timing of the report, the organizational support, the "ownership" of the study, the evaluators' roles vis a vis policymaking. These elements combined to create a salubrious climate for evaluation utilization. The unique elements can be summarized as follows:

1. The study focused on one question, a question which legislators, SEA managers, and school people needed answers to. The call for the study contained

no laundry list mix of questions, no prescription of methodology, no assumptions about relevant variables.

2. The evaluators themselves owned the evaluation. The evaluation team conceptualized, designed, negotiated and defended their decisions:

-- they thrashed out a case study design which allowed ECE processes to emerge from observation and analysis. Their team brought perspectives from various units which affected ECE.

-- their methodology was a systematic version of the "methodology" that had been used by legislators and SEA people--observing schools and listening to client perceptions.

-- they maintained high level support. They remained attuned to political and organizational tensions.

-- they described processes and problems in narrative style with "implications" for policy.

-- they were a team formed for and by this project so a team was committed to its success.

-- the team worked on just this project, from design to dissemination.

-- they used the findings in intensive discussions in SEA policy meetings.

-- they used the evaluation in the field.

These are important elements, as will be seen from the summary of ECE evaluation history which follows. The reader will see ECE as a policy, as a program, as a political event, and note the ensuing evaluation dilemmas.

The history of ECE evaluation and policymaking was complicated by its political nature. (In an earlier report (Marshall, 1981) the author described the state policy system and its effects on evaluations and decisionmaking for ECE.) ECE evaluations had been fraught with difficulties due to political

tensions, ambiguity in program goals, overly ambitious and conflicting program goals, disagreements on appropriate methodologies, criteria and on the validity, objectivity, and utility of previous evaluations.

ECE evaluation dilemmas followed from the fact that it was a political program and with various and changing goals. What was ECE? The answer depended on who you asked and when. State Superintendent Wilson Riles had made election campaign promises about reforming and revitalizing education, raising reading and math achievement, and focusing on very young children. His "Blue Ribbon Commission" of experts and concerned citizens had made recommendations about extending schooling to four-year olds, addressing socioemotional needs of children, individualizing instruction, developing children's self-concepts. During policy formulation within the SEA with school people and during negotiation with the legislature, ECE was transformed to include goals for restructuring education systems with school-site decisionmaking and School Advisory Committees. Legislators' concerns about controlling the purse strings, raising achievement scores in low achieving schools, and maintaining legislative control over the SEA would be met through yearly evaluations. During the SEA's operationalization of ECE through guidelines, application forms, state Monitor and Review (MAR) processes, and through SEA technical assistance processes, ECE took on elements of state-controlled reform to meet state-determined goals. During locals' implementation of ECE, its implications for changing roles and structures in education systems began to show.

The mixing, conflicting, ambiguous, and political goals enmeshed in ECE created the following evaluation dilemmas:

1. who should evaluate? with what resources, personnel, and methodologies?
2. which goals should be measured?

3. how could one measure such goals as quality of parent participation or self-concept development?

4. was controlling SEA the main purpose of evaluation?

5. was program improvement the main purpose of evaluation?

6. was evaluation for state-level or local-level use?

7. if evaluation was for cost-effectiveness, what constituted effectiveness?

8. how much time was reasonable to allow for implementation of ECE before results would show?

9. could ECE schools be compared with non-ECE schools, given that lowest achieving schools were the first ECE schools?

10. could ECE effects be compared with other programs, such as the Miller-Unruh Reading Programs, given that ECE goals went beyond reading goals?

11. what evaluation questions were pertinent?

12. would not ECE implementation and morale of participants be adversely affected by a critical evaluation?

13. what aggregate state-level evaluation design could take into account the fact that school sites determined their own goals?

Initially these dilemmas were managed by standard operating procedure: the SEA evaluated ECE by aggregating measures wherever they could be obtained and by printing a pamphlet, Early Childhood Education: First Annual Evaluation Report, 1974-75 (1975). However, the report was compiled in a hurry with data from schools which had just begun to implement ECE. Riles, in his introduction to the report, proclaimed that ECE "resulted in significant improvement in school achievement" (1975).

When Riles requested funds for ECE expansion in 1975, legislators, the

Legislative Analyst, and the Department of Finance, the media, and several academicians zeroed in on the evaluation. The critics of the evaluation noted the following problems:

1. only 43% of ECE students were represented in the data, and the data were generally inadequate for supporting the claims of achievement gains,
2. no evidence was presented on the goal of parent participation, and other ECE goals,
3. the SEA had failed to abide by the law because of its inadequate evaluation,
4. there was no comparison data,
5. there was no control for the Hawthorne effect,
6. there was no control for the mixing of ECE funds with the funds and programs of other state and federal efforts at school reform, compensatory education, and reading,
7. there was no effort to link achievement gains to ECE processes,
8. there was no effort to verify school district data,
9. there were no descriptions of how ECE works,
10. it was inappropriate for the SEA to evaluate its own programs.

Legislators were confused by the evaluation dilemmas. They wanted to know how the program worked and whether or not to put in more dollars. One senator who attacked ECE was, within an hour, deluged by phone calls from parents and school people who supported ECE. ECE had its own constituency and legislators saw that Riles was emotionally and politically committed to seeing ECE expand; that Riles was a charismatic elected official who could build a constituency through dispensing state dollars.

The legislature approved ECE expansion but mandated an independent

evaluation of ECE. The SEA, Department of Finance, and legislative staff contracted the independent evaluation with the UCLA Center for the Study of Evaluation (CSE) and negotiated evaluation questions, design, measures. This CSE evaluation had to include control groups, norm-referenced tests, and it had to be attuned to the concerns of the contractors; policy groups demanded evaluation information on the MAR process, on SEA management of information, on parent involvement, staff development, and on instruction variables such as the diagnostic/prescriptive mode and individualization.

These information demands, the limited funding of $200,000, the delays in transporting achievement score data to CSE, pressure from the SEA to send an Executive Summary before the evaluation was completed, suspicions due to the friendly relations between the SEA and the CSE, and other difficulties complicated the CSE evaluation. When the premature Executive Summary and the four volume evaluation report did not coincide completely, legislators attacked the evaluation. Further, legislators still could find no description of how ECE worked. When the CSE director reported to legislators, she reiterated her researcher cautions about drawing inferences, repeating the report's statement that "our study cannot provide unequivocal guidance for legislative or school decisionmaking because evaluation and methodological concerns were subordinated by program implementation decisions" (Baker, 1976, p. 9). Although the evaluation did point out MAR and staff development problems that led to new legislation, it still left legislators angry that evaluations could not, or would not, make clear, definitive policy recommendations for them.

Thus, by 1977 ECE evaluations had been attacked and politically manipulated; program managers were not getting guidance from evaluation, and evaluators were confused. Legislators were disillusioned by evaluation reports of questionable

credibility, as when SEA evaluations seemed to be promotion documents used to lobby for expansion. They were also frustrated by huge evaluations replete with statistics but lacking program descriptions and policy recommendations.

SEA reactions to negative findings had been close to censorship, or at least obfuscation in the 1973-76 period. SEA reactions to criticism had included belligerent attacks on the credibility and expertise of the critics and the marshalling of the ECE troops. But in early 1977, Riles began to show a new openness--a willingness to face implementation problems and seek solutions and a willingness to look at goals other than achievement. Teacher unions, key senators, and legislative advocates for administrators, school boards, and the large cities were actively voicing complaints about ECE. The most charismatic state superintendent could not manipulate such an audience into ignoring San Diego's own evaluation of ECE which showed that ECE actually had some negative effects. There was no attempt to downplay the findings in the Evaluation of ECE, ESEA Title I and EDY, 1975-76, even though it showed negative effects of ECE. SEA was taking a new stance toward ECE and toward evaluation, as shown in this quote:

> "When we noted the decline, we generated possible reasons and sent it to [an SEA official], who said 'here's a problem, we should be honest about it and then go out and find the problem source.' There was no attempt to suppress the finding." (3/28/80)[2]

The SEA realized that it must look at the relations between its processes and change processes in schools. It saw that regression analyses on selected variables and achievement scores did not inform SEA as program managers. The stage was set for a new type of evaluation. It was an opportune time to answer questions like, what research methodologies are preferred? what are

the crucial timing, reporting presentation modes, and policy-relevant evaluation questions which will facilitate evaluation utilization? and, what strategies will enable a new evaluation attempt to survive organizational and political tensions? Policymakers and program managers were searching for answers which previous evaluations had not provided. The SEA began The Interim Report evaluation at this crucial window when both policy makers and program managers agreed that they needed answers to one focused question.

This ECE evaluation history background shows that ECE evaluations were replete with problems in meeting the criteria for good evaluations. The following list shows that ECE evaluation had problems in:

1. internal validity: no comprehensive representation of ECE was reported,
2. external validity: whether or not the findings were generalizeable to other situations was unknown since the linkages between program and findings were unclear,
3. reliability: accuracy and consistency were questioned in each evaluation; questions arose about how to compare different test results, how to separate ECE effects from other variables,
4. objectivity: questions arose concerning the representativeness of samples, adequacy of data collection techniques,
5. relevance: the emphasis on achievement scores missed substantive restructuring goals of ECE,
6. importance: policymakers and program managers needed to answer "does ECE work? should it be modified? continued? expanded?" The evaluations failed to provide answers,
7. scope: focusing on achievement scores and counting numbers of participants did not capture the scope of ECE, a program aimed at reforming and re-

vitalizing education through changes in governance and instruction,

8. credibility: by miscalculating the degree of legislative scrutiny, SEA evaluations lost credibility; the CSE evaluations lost credibility from the Executive Summary mixup. More importantly, it lost credibility because it was replete with caveats which policymakers interpreted as weakness,

9. timeliness: although the evaluations were reported in legislative committees before funding decisions were made, the real information-gathering and planning by program managers (SEA and ECE schools) had to be carried out far in advance. SEA budget proposals were made long before evaluations were reported. SEA information gathering and alterations of policy occurred far in advance of the reports,

10. pervasiveness: the SEA evaluations were disseminated to more than their intended audience. SEA did not anticipate that legislative scrutiny would be buttressed by critiques from academics. The SCE Executive Summary mixup was from poor dissemination planning,

11. efficiency: neither policymakers nor program managers got maximum benefit from the evaluations due to political tensions between SEA and legislators. Neither OPER nor CSE had budgets and time for a maximal search for evaluations questions and for researching the processes and designing their measures.

ECE evaluations had been a series of attempts to answer the wrong questions and mismatches between evaluators' agendas and audience agendas. Any new approach would improve the situation.

THE INTERIM REPORT

Windows for Evaluations

The Interim Report was conducted at a time when alignment of interests of

policy and program people and available evaluations resources allowed a useful evaluation to develop. The climate in California educational policymaking skies was open and ready for answers. SEA defensiveness about ECE had toned down. Legislators' need for information had increased since each previous evaluation had merely complicated decisionmaking and failed to describe ECE implementation and processes. The favorable alignment of variables in the policy arena supported The Interim Report. The evaluation could not have developed nor could it have been useful if the policy variables were not so aligned. Evaluations, like space shots, need maximum feasible windows.

The frustrations with the prior ECE evaluations set the stage. The intricacies of The Interim Study story which follow provide insight for identifying policy arena receptivity to evaluation.

Evaluation Basics: Evaluation managers must face these fundamental questions:

1. what is the precise policy or management question?
2. who is the audience?
3. what purpose will the evaluation serve?
4. what research design is appropriate for the question?
5. what are the biases, intrigues of politics, and constraints which affect the potential audience?
6. what human and monetary resources are essential for the study?
7. how much should evaluators reach into policymaking--should they take forceful stands to ensure that their findings be incorporated in policymaking?
8. how should the findings be reported?

Such a list may outrage evaluation purists--the idea of letting themselves be affected by political intrigues would be horrible. Nonetheless, the story of

this "successful" evaluation suggests that evaluation managers must face these questions.

The Interim Report evaluators searched for answers to these basics. The answers and the stances they chose enabled them to produce a useful evaluation. The outline of evaluation basics serves as a framework for recounting and analyzing The Interim Report study.

What is the precise question?

A focused question to frame an evaluation emerged directly from policy deliberations. The Evaluation Report of ECE, ESEA Title I, and EDY, 1975-76 reported that:

> "In schools whose entering students averaged between the 21st and the 99th percentile on the 1973-74 Entry Level Test, grade three reading achievement imporoved markedly beyond predicted levels after three years in ECE . . . . In schools whose entering students averaged below the 20th percentile on the E.L.T., grade three reading achievement declined relative to prediction after three years in ECE." (SEA, 1977, p. 36)

In reporting this finding to the Joint Legislative Budget Committee, to Senate Finance, Assembly Ways and Means, and to the education policy committees in the Senate and Assembly, the SEA made a commitment to explore the reasons for this finding. Senate Finance effectively mandated the study in Item 283 of the Supplemental Language to the 1977-78 Budget Act. The language entrusted to the SEA the task of selecting the appropriate unit, personnel, and design to answer the question, but demanded answers to one focused question: "What are the characteristics of the low achieving schools and what processes are causing the decline?" No demands were added; no requirements of comparisons with other

programs; no assumptions as to causes were built into the question and no evaluation measure or methodology was specified.

What purposes will the evaluation serve? What are the biases, political intrigues and constraints which affect the potential audience? Who is the audience?

The audience included the SEA as program manager and the SEA as a political unit. It included interest groups, legislators, county and LEA program managers and ECE participants and the media. (See Figure One.) Each of these audiences had their own agenda which an ECE evaluation could potentially serve. The biases, intrigues and constraints are illustrated in the following descriptions of the audience.

After four years of ECE and no descriptive, definitive, credible, comprehensive evaluation, information needs were high. The various evaluation audiences had raised questions and complaints. Some expressed outrage for low scores on the state Monitor and Review and contented that the SEA was inflicting a "one best system" approach through its guidelines. United Teachers of Los Angeles hired a truck to carry to the legislature the reams of forms which they saw as a symbol of overcontrol, overtesting and burdens from programs like ECE. Senator Jerry Smith had criticized the SEA MAR teams and technical assistance and told Riles to keep SEA people out of his district. (When Riles replied that every school in California was his district, this did nothing to defuse the political nature of any information on ECE implementation and success.) This audience wanted to ferret out the mismanagement and political control they saw in ECE.

For the legislators as part of the evaluation audience, the 1973-77 evaluations built a record of mistrust of machinations-by-evaluations and of Riles'

The ECE Evaluation Audience

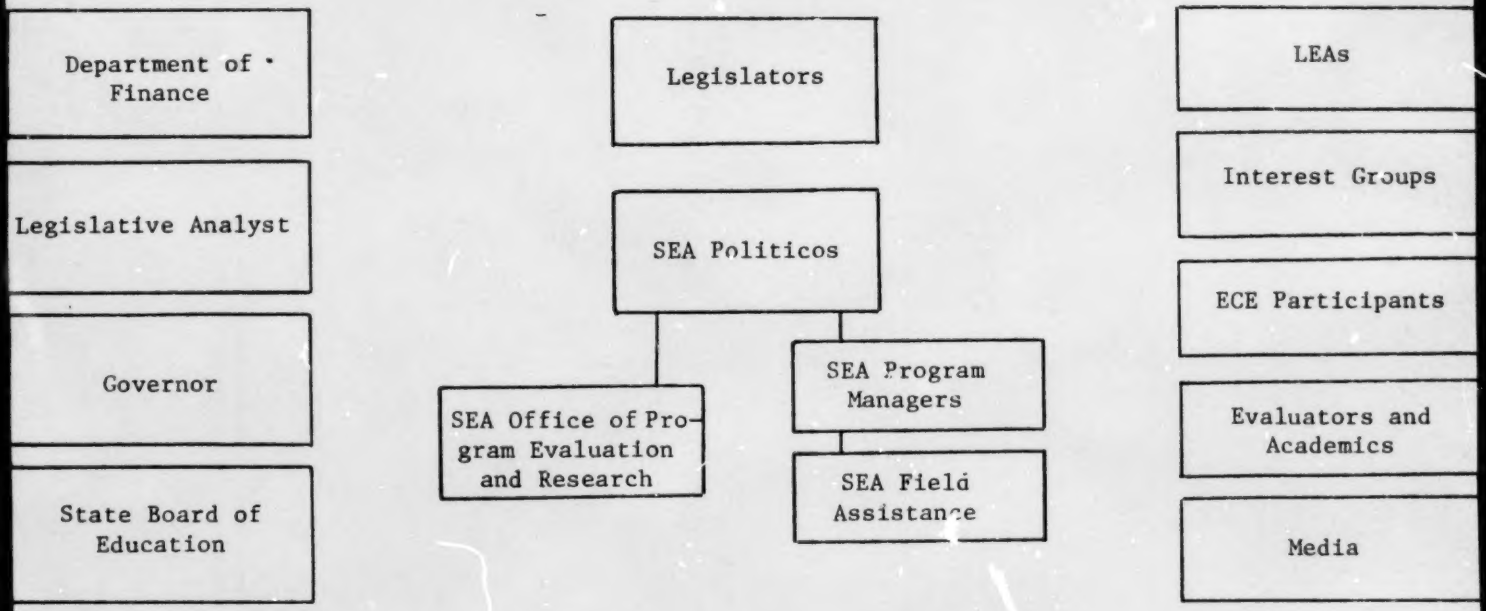

FIGURE ONE

program promotion through exaggerating favorable findings and discrediting critics. Any SEA evaluation had to confront and rise above this suspicion. In addition, legislators were frustrated by long reports which detailed measurement procedures but failed to describe how the program worked. Legislators wanted information that could systematically describe and substantiate the successes and problems with ECE; this would enable them to alter educational policy, answer constituents' complaints, and alter resource allocations to maximize effectiveness. Also, the legislative audience wanted evaluations to help them keep their control of education policy.

The audience which consisted of SEA politicos, program managers, LEAs, ECE participants and SEA field assistance people needed information to make ECE function smoothly. In the past, SEA organizational tensions and politics had complicated the conduct, reporting, and the utilization of evaluation. (Figure Two shows a functional chart of SEA units involved in ECE.) There were tensions between SEA program managers and OPER:

> "Evaluators <u>are</u> defeinsive about their purity. But evaluators <u>don't</u> understand the program so they can't measure it or describe it appropriately. OPER's numbers don't do it. They were always off on their description. Their numbers were fine but their conception was off--it was not in the flavor of ECE." (9/5/79)

In addition, program managers relied more on the information network with LEAs and participants; it seemed more relevant than the standard deviations and interaction analyses in evaluations. More often than not, program managers had changed policies long before an evaluation report came up with a finding which said that policy should be changed. The information network did not run smoothly, however. The dissemination of information on policy

SEA Units Involved in ECE Evaluation, 1976 - 1977

(organization chart of functions and key people)

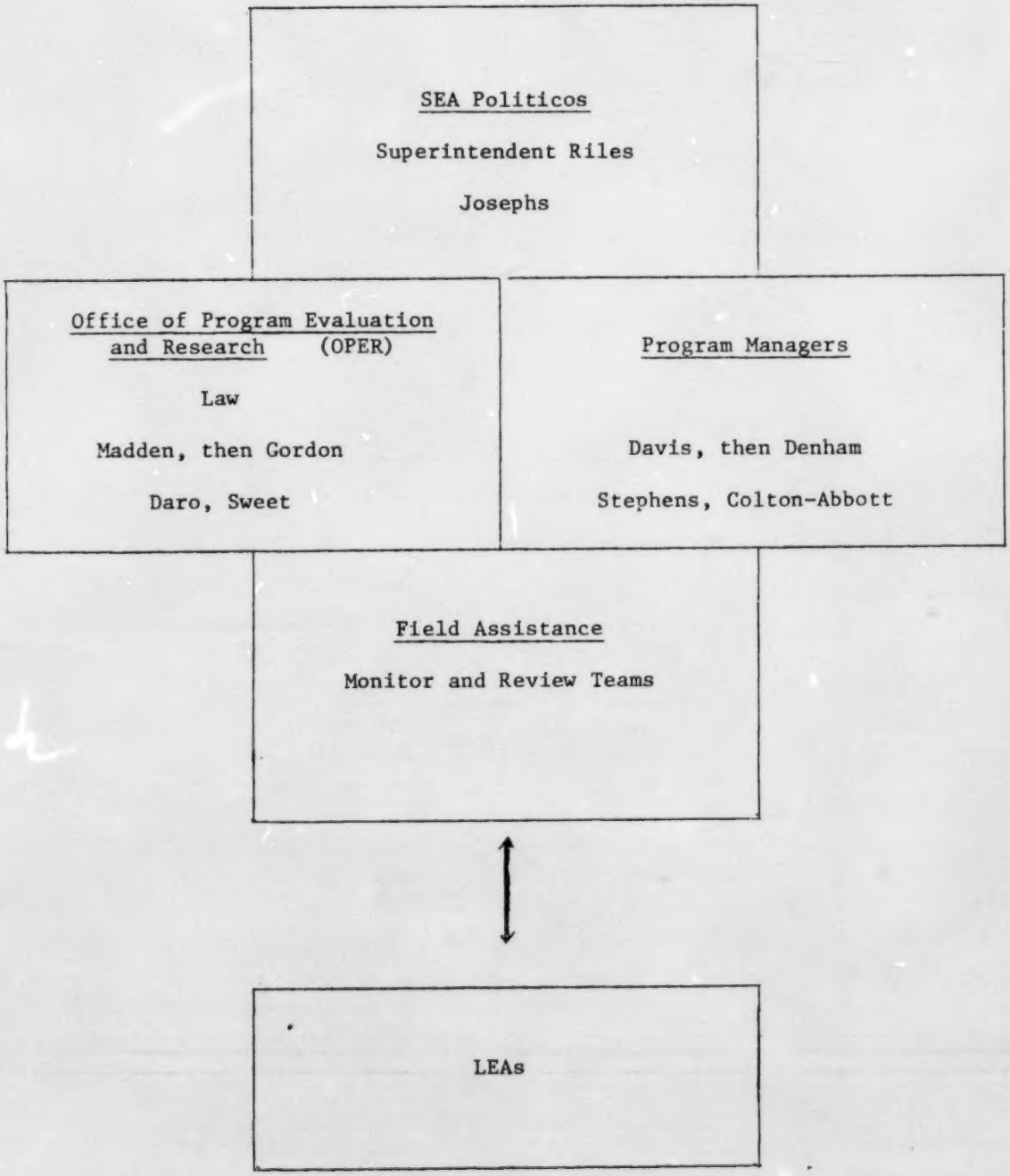

FIGURE TWO

changes, changes in forms and guidelines was irregular, and led to awkward situations, as a county ECE director described:

> "Our role at county is to continue being supportive of SEA when they've blown it. It is difficult, like when they send letters promising documents and then don't come through, and when state people gossip so they leak the fact that a requirement will be phased out before it is official and I'm there inservicing people on how to fulfill that requirement." (5/22/79)

When SEA field assistance teams went on the road they faced school people who needed translations, or at least interpretations, of unclear SEA guidelines. They faced unbelieving school administrators who recalled having followed an SEA policy only to find that the policy interpretation from one SEA person differed from that of another SEA person. Policy decisions made by SEA's politicos did not reach the field assistance people immediately. In addition, the field assistance and MAR personnel, the SEA representatives who were most visible in the field, were looked upon as the least competent SEA employees.[3]

Thus, the program manager audience needed evaluations that would see ECE as a whole program and look at implementation issues. Their present mode of keeping close contact with school people had helped avert drastic problems in implementation, but they needed more systematic, systemic, and more timely program evaluation.

Tensions between OPER and SEA politicos had complicated SEA evaluation utilization. Although the OPER director, Law, called it S.O.P., others said reports are changed:

> "If Josephs says no, it won't go. She'll say it doesn't have the right tone; it doesn't tell how the program works; it's not positive

enough." (9/5/79)

The practice of sending all evaluation reports to the SEA Politicos for editing did cause consternation. OPER took pride in doing objectively scientific studies. In 1976, the second in command in OPER, Madden, was replaced by Gordon who was seen as politically astute, tied in with Program, not a numbers man:

> "Gordon is slick, cool, really going somewhere . . . . He was part of the Executive Cabinet and that says you're somebody." (8/31/79)
>
> "Gordon was trusted by both sides, Program and OPER. Madden was more the professional evaluator; he was not policy-minded, didn't play games . . . . Gordon sees the bigger picture of what evaluations have to do." (9/6/79)

This personnel change occurred around the time when the legislature was considering elimination of OPER. OPER took on a new perspective in late 1976, as illustrated in this quote:

> "I used to be rigid, a pure evaluator, upset about Program managers altering our descriptions. Now I've seen what can happen to a good program. Senator Carpenter uses evaluations against us when evaluations show a real problem. Evaluation should be to identify problem areas, to give honest answers. They can't when they're used to destroy a good program." (9/6/79)

Thus, tensions existed among the functional units within the SEA, and these tensions were exacerbated by SEA relations with LEAs, interest groups, and the legislature.

Audience includes the press when a controversial program is evaluated. The Los Angeles Times' reporting had mobilized critiques of SEA evaluations. Riles

had used television spots to publicly interpret the finding of the CSE evaluations. The press would likely publicize evaluation findings again, especially achievement gains and losses. The SEA politicos would have to answer to the public and to the legislature. LEAs would have to answer questions and concerns of their clientele, especially if negative findings were reported.

Finally, the evaluation could have a new audience consisting of whatever groups were involved in the formulation of major new educational policy. During spring and summer of 1977, the SEA began preparing policy proposals which eventually became Assembly Bill 65, the reformulation of ECE and the formulation of a program to improve intermediate and secondary schools. Information about how schools implement ECE--a state-directed, locally planned restructuring--could be valuable for this policymaking agenda.

Any 1977 ECE evaluation could be attacked, supported, used, or abused by these audiences depending on the evaluation management.

The Interim Report evaluation could serve legislators if attuned to their biases and frustrations. It could serve program managers if it got above the details of operations and looked at the whole process of ECE. The evaluation could serve the public, through their legislators and through the media, by raising questions about schools. The evaluation could inform vital decision-making if it could feed into the formulation of A. B. 65.

What research design is appropriate? What management, human, and monetary resources are available?

Decisions on evaluation design and resources for The Interim Report deviated from past practice. When Senate Finance requested an explanation for the declining scores, the SEA program manager, Davis, took responsibility. However he sought input of two OPER people, Daro and Sweet, and Denham, in his

own unit. Since Senate Finance had not appropriated funds for this study, person hours and money had to come out of existing resources. The usual pattern would be to give the study to OPER evaluators, while SEA program people continued their regular activities--wait, then read the findings, and reject them as non-representative of the real program. The Interim Report deviated from this pattern. The management, personnel, design, and reporting of this evaluation were intertwined and unique arrangements.

Management:

When Gordon moved to OPER he maintained his connections with Program and Policy people, especially Josephs, and built bridges with the legislature. OPER management and connections were shifting.

Program management shifted during early 1977 also. The management style of the Director, Davis, has been appropriate for operationalizing an amorphous reform program, assertively creating guidelines and structures, defending and promoting ECE. Now, SEA policy had shifted to a stance which allowed controlled criticism and increased program flexibility. Davis was going, whether he did not adapt or whether he was a sacrificial lamb, to telegraph this policy shift to legislators and school people, he was going. His leaving was tied into Interim Report management, personnel, and design.

As Davis and Denham brought others into the discussion of the evaluation, their control dissolved. A unique group was forming: Daro in OPER and Stephens, in Program added their special perspectives. Daro expressed his orientation to evaluation as "not fixed on measuring objectives." (3/29/80) Although Daro was in OPER he was readying to break away from traditional evaluation modes. In addition, he was one of Marian Josephs' valued advisors as observers recounted: He "had Marian's ear." (3/26/80) "Marian never reads anything but she relies

heavily on what a few trusted individuals tell her." (4/29/80) Stephens' work included a case study of an innovation which showed the importance of building team commitment to a project by team involvement in planning. She was new in SEA program and recalled:

> "I had a lot of energy then and I didn't want to just do a schlocky job. I'm not sure that now I could do it, having been socialized. I wouldn't have yelled as loud or fought as many battles." (3/27/80)

Stephens was ignorant of SEA organizational patterns. She was not constrained by the antagonism between OPER and Program and she was not socialized to keep her ideas quiet or allow SEA politicos to moderate them.

Daro, Stephens and several other OPER and program staff coalesced and, with tolerance from program managers and encouragement from Gordon, began the task of arguing an evaluation design into existence. Several recalled the process:

> "The evaluators on the team assumed a whitewash would come out of it. We had emotional stuff to work through. It went from a stormy to a cohesive group process." (3/28/80)

Management of the evaluation was evolving to be non-management. The team took control of itself. Fundamental issues came to a head at a meeting in Board Room with team members--Gordon, Davis, Denham, Josephs, and other policy people. Riles was travelling but still involved. The setting and the actors indicated high level decisionmaking. The issues on the surface were the following: (1) will the evaluation be strictly research (as the team wanted) or will it be combined with field assistance (as Davis wanted)? (2) will program or OPER manage the evaluation? The underlying issues were: (1) how

can Davis be undermined? (2) how much high level support exists for this evaluation? If it has negative findings which make SEA look bad, who will defend the study?

Daro recalled the meeting:

> "Marian [Josephs] set it up ahead of time. She phoned Wilson [Riles] and told Glenn [Davis] and me to talk to him to resolve the clash over the study. It was 40% staged." (3/26/80)

The meeting ended with Davis' ideas defeated; with Gordon managing, promoting and supporting the team's case study design but with the understanding that this unique team would essentially manage itself. It was also understood that Denham would manage Program's side of the evaluation but would essentially let it happen and that there was a commitment of strong support from Josephs and from Riles.

It was clear that the high level support had to be nurtured. That became Daro's job because of this access to Josephs. It would also be Gordon's job because of his political skill and good relations with program and politicos and because he was willing to promote this case study approach.

The unique self-management, cross-department staffing and high level support was settled, along with the commitment to the case study methodology.

Design and conduct of the evaluation: The team had hammered out the design through arguing and through training sessions, but their approach was not assured until the Board Room treaty described above. The team argued itself to consensus on these issues:

1. case studies could provide answers and could withstand questions of limits of generalizeability, validity and reliability,
2. the case study approach was expensive in person-hours and required special

expertise but the high level support would allow for it,

3. the team should be the designers, evaluators, and reporters of the evaluation and this type of evaluation must be explained, not just reported.

Daro described the considerations of the case study approach:

> "It was high-risk, high gain. If we failed, it was a reflection on us, not on the organization. We were not a part of the main process, nor was the study . . . . We saw case study as appropriate for explanation, not just for description. Not "what," but "why and how." The old way made inferences for "what" without the whys and hows. That's not effective." (3/28/80)

Stephens' ideas concurred:

> "It's from the triangulation approach in evaluation. It's another way of gathering data. In the past, we'd put all our eggs in the basket of stats. Like the first CSE study had chart after chart of variables saying what happens if you cross parent involvement with curriculum planning, etc., then making generalizations based on these variables. This [case study] seemed a way of combining statistical data from test scores with more qualitative data and putting in some human judgements." (3/27/80)

Although the evaluation was designed by group input, Stephens and Daro took leadership. One OPER staff person pulled out, rather than compromise his training:

> "He felt he couldn't be part of anything that was called evaluation but was so unstructured. I kept saying it isn't unstructured but structured a different way. If you didn't have an instrument he wouldn't be part of it." (3/27/80)

High level support for alternative evaluation approach did exist. Denham re-

called:

> "We had Josephs, Davis, me and others in a meeting with Law about using achievement data. We talked about the utility of summative evaluation--how it doesn't help Program people nor schools, it's demoralizing. It does no good to evaluate bad schools a third time. Instead we need to help them learn something about the nature of ECE. The theme was 'do something different.' (3/28/80)

Law said he knew the value of case study evaluation but was hampered because legislators had expectations that experimental studies would supply them with definitive information.[4]

The commitment of staff resources required the case studies makes them expensive. The design, observation and analysis, the need for a team of experts pulled from several units meant that high level support across SEA units was crucial, since the case study approach was expensive and findings might not be generalizeable. The main selling point was that it was the most appropriate, given the need for explanations of ECE processes.

The sample selection issue raised basic problems. Politicos wanted larger sample, criteria for selecting "increaser" and "decreaser" school were controversial; Program wanted to provide technical assistance mixed with evaluation and wanted Oakland and Los Angeles to be included. The team needed to maintain Program and Politico support to get access to schools. The design maintained controls by the team training and its structure. The variety of expertise within the group built in a tensions which was a control for validity and reliability. The in-group tension was modified by a developing pride and defensiveness about the study:

> "It was hard fought to do the study originally so the commitment was strong. We had a collective defensive attitude . . ." (4/29/80)

"There was resentment toward the 5th Floor [Politicos]. We were assuming that our efforts would be quashed or washed." (3/28/80)

The struggle to obtain and maintain high level support for the study created team energy and determination to do a piece of significant work:

"We were thinking we were prima donnas even if nobody else was and we wanted to do something that was useful. Otherwise, why do it?" (3/27/80)

This sense of pride and "we'll show them" attitude was an informal control. The structure of school visits built in controls. Two team members spent two days looking for what, why and how the school was a decreaser or an increaser. They would each form hypotheses about processes in that school, devise techniques and data collection schemes to test those hypotheses, then return to the school one week later to "test" the hypotheses. The team members were a quality control and bias check on each other.

Team members went to the schools with their observation skills but without assumptions. They looked for processes that contributed to decreases or increases for two days. They made independent hypotheses, discussed them during the intervening week, devised appropriate tests, then spent the final two days checking hypotheses. Each team wrote a case report for each school describing the processes identified and predicting whether or not the school would change its achievement score pattern.

What should be the report format? How involved should the evaluators be in decisionmaking? After months of planning, training, and data gathering, The Interim Report team faced reporting issues. These issues were intertwined, and the intense team commitment propelled their resolution. A report of all cases would be too long, would breach confidentiality, and would make no

recommendations.

A writing team analyzed the cases looking for patterns and writing policy recommendations. Meanwhile Daro was continuously informing the politicos of the progress, format, and findings. Politicos pressured to include a section on statistics. A member recalled:

> "That's Chapter II. The statistics part was weak . . . . It says nothing. We all hedged about putting it in and when we saw it we said, 'oh, barf.' The study stands alone, without the stats." (1/16/80)

Daro "scrounged around," spent all night with computer printouts and put together statistics on background variables.

Tension over reporting findings had existed from the beginning. The team was determined to ensure that their work was used; the initial high level support and Daro's continuous communication with politicos set the stage. The question remained, however, about whether politicos would allow dissemination of findings that implied poor communication to schools and inadequate implementation assistance from SEA or findings that showed that a major component of ECE guidelines--the diagnostic-prescriptive approach to improve instruction--had contributed to mechanistic teaching, more paperwork and less teacher-pupil interaction. SEA evaluation reports had never been so blunt about shortcomings, especially a program which was called "Riles' baby."

The writing team wanted to be hard-hitting and frank yet they worked with the realization that politicos could modify their writing. One member recalled:

> "We had continuing debate about how strong the statements could be. It came down to how strong the person felt it was essential. This was tempered by SEA norms . . . . These implications statements were probably the strongest statements ever to come out of an SEA written report." (4/29/80)

Through negotiations with the Office of Government Affairs, a compromise format emerged. Chapter II would satisfy the objections of the statistical types. The report would give the rationale, strengths, weaknesses, and limitations of case studies, including a statement that case studies have "less control over the emergence of an incorrect explanation." (State Department of Education, 1977, p. 3) The final report uses six of its 35 pages to explain the case study approach.[5] This methodology was new to many in the evaluation audience and the team was determined to ensure that their findings were considered valid and significant. Chapter III included the findings which showed inadequate communication of the intent of ECE, the need to recognize the importance of leadership, teachers misinterpreting the ECE goal of individualization and lowering expectations for low performing students, poor planning for staff development and inadequate inservice for improving staff capacities to serve ethnic and language minorities. The findings showed that some teachers interpreted the diagnostic/prescriptive mode to mean self-teaching with commercially prepared programs, that new ECE funded reading programs were not integrated into the school curriculum, that school people were confused by the various evaluative techniques, so that the inadequate attention to implementation processes were effectively thwarting progress toward ECE goals. After each section on findings, the team placed an "Implications" section. These were policy recommendations--hard-hitting ones. For example:

--Inservice training should help adapt new programs and skills to the preexisting practices rather than merely enforcing compliance to a new program. (p. 25)

--Within the reading curriculum itself, a balanced variety of reading skills should be taught. Scores declined when there was overemphasis

on decoding and phonetic analysis and underemphasis on integrated skills such as comprehension and story writing. (p. 30)

--The strategic implementation parts of ECE--planning, implementation, internal and external program quality review (including MAR), and evaluation--need to be more explicitly connected to each other. The ongoing interaction among these parts has been overlooked in some schools. (p. 32)

Such statements clearly and concisely stated fundamental problems with the initial planning and guidelines and field assistance. Politicos allowed such a report although SEA staff morale and ECE participants could be hurt. Politicos were open to the critical findings and policy implications because of Chapter IV entitled "Department Action" which gave SEA the change to explain its proactive "ongoing effort to seek appropriate ways of improving departmental, district, and school programs."[6] (SEA, 1977, p. 33)

By the time The Interim Report was printed, the SEA had undertaken intensive reexamination of ECE policies in preparation for A. B. 65. Chapter IV outlined these efforts and emphasized that The Interim Report study design did not allow for estimates of how many ECE schools were having the observed problems. The chapter answered the anticipated questions from riled-up legislators and school people who would demand that SEA do something about the problems. It offered promises of improvement which could convince doubting legislators and school people that the program should be salvaged, despite the problems. The report-writing controversies then were settled.

Questions remained about the evaluators' role in reporting and disseminating findings. They were determined to ensure that their work was significant. The timing was perfect.

Decisionmaking was a-brewing and there were indicators of department openness to reexamination and reformulation at the beginning of the evaluation (see p. 11), and in the summer of 1977, Josephs structured "half-day sessions" for that purpose. Participants included people from the field assistance unit, Program, OPER and politicos. It was another ad hoc, special purpose unit. One participant explained:

> "The timing was great--we finished [The Interim Report] in June and A. B. 65 was coming out. Marian [Josephs] was wanting to get our level of people--the people who are the consultants and the people who are in the districts all the time--together and say, 'now we've got a chance to rethink and redo ECE.' We had a whole week of half-day sessions at a hotel. She wasn't purposefully pulling together the peoole who were on the team but those of us who were there were loud enough that the stuff from the study became the stuff that laid the groundwork for getting A. B. 65 rebuilt." (3/27/80)

The half-day sessions coincided with the report writing. Daro recalled:

> "We were pulling together and using our findings at the half-day sessions . . . . We talked about the study to everyone, in every form we could. I'm sure it was proclaimed in more emotional ways than it was in print." (3/28/80)

Denhan recalled the sessions:

> "Like pitched battles at first but later more collegial. Under Marian's leadership they were at a high conceptual and integrative level, yet they were very concrete about what goes on in schools." (3/28/80)

The evaluation findings fed directly into this unique structure for intense decisionmaking through perfect timing, through Daro's access to Josephs, and

because team members were determined to see that their work had impact. In addition, this decisionmaking required information on process and implementation. The Interim Report findings were providing this essential information in depth and in personal reporting.

Clearly, no distinction was made between designer, evaluator, evaluation reporter, and decisionmaker. The team planned the design, pushed it through for approval, carried out the study, vociferously reported the findings, cultivated relationships with politicos to ensure that the findings would be used, wrote the report with clear policy implications that could not be soft-pedaled, and then actively proclaimed and promoted their findings and the implications as participants in a key decisionmaking body. Stephens had planned for this:

> "One of the things that was so important to how it turned out was that the group that were the deliverers of it were the doers of it . . . . What they typically do is get two stars to design something and then they train people for three hours and then do it . . . . We fought to pick who we wanted on the team. A key thing is ownership. We wanted their time and the only way to have the team own it was to have them there on the planning . . . ." (3/27/80)

The written report progressed slowly. It finally went to the legislature as an "interim report" in Fall 1977, and later the findings were incorporated in the yearly Consolidated Application Evaluation Report (State Department of Education, 1978). Team members heard indirectly that it was well-received by legislators and legislative staff.[7] The SEA printed one thousand copies and the supply was soon exhausted. Stephens personally disseminated the report to school people; Daro used the report in SEA workshops and at the California

Evaluators Research Association to illustrate an approach to evaluation which looked for effects and for explanations of why schools differed.

The evaluators' personal commitment to the process and the product of The Interim Report study guaranteed that their audiences would take notice. Daro believed in the appropriateness of evaluation to describe and explain processes in schools and Stephens believed that ownership of one's work ensured high quality work. Both put their beliefs into action and wanted to let the world know. The final version of the entire study did not appear until March 1980, a full two and a half years after completion of The Interim Report.[8]

This account of Interim Report elements details crucial strategies for managing the very real organizational/political diplomacy while conducting a credible, useable evaluation. Evaluations which are utilized for program improvement and policymaking are unique. Such evaluations merit attention and analysis. The Interim Report, which used a case study methodology on a comparitively small sample--had minimal editing, lacked a glossy cover, included few numbers and charts and, most of all, was used in policymaking--deserves analysis. Its approach allowed the emergence of questions and answers on ECE processes. It was the first evaluation that looked at ECE as a restructuring of the entire educational system for K-3. Although original ECE goals demanded restructuring of systems and governance, previous evaluations had used methodologies to measure variables and test for significance but failed to look at processes and system. Too, previous evaluations had been used as ammunition in the SEA-legislative battles, so their usefulness for decisionmaking was diminished. The Interim Report evaluation designed, carried out, and reported an innovative approach to evaluation, avoided the pitfalls of previous evaluations, and its findings were widely used for decisionmaking. The details of

the management of this evaluation should provide clues that can guide evaluators.

Analysis of Successful Evaluation

The description of the history of ECE evaluation and the constraints presented by ECE evaluation audiences provides background of this successful evaluation. In spite of this history, The Interim Report, an un-glossy, minimally edited, methodologically innovative evaluation, was received as credible by legislators and SEA policy people alike. A first! The SEA used this success to substantiate the value of orienting OPER to more small, focused studies which used a variety of exploratory methodologies. The SEA incorporated the recommendations of The Interim Report in its mode of relating and in the messages relayed to counties, districts, and schools. SEA Program and Field Assistance staff used The Interim Report in workshops. It was a tool for training them to look at processes of implementation as well as a document with possible improvement strategies. Requests for copies of the report came from many school people in other areas, so this approach had spill-over effects. The study was a part of the training for Program Reviewers. (MAR was renamed Program Review.)

Another practical application came from the fact that The Interim Report was well-received in the legislature and in the SEA, lending credence to earlier suggestions that evaluation should focus on processes of ECE implementation.

Most importantly, The Interim Report findings were incorporated in legislative and in SEA policy program planning and field assistance.[9]

The intermixing of key elements which led to successful utilization are

illustrated in Figure Three. In this case, evaluation utilization had nothing to do with nicely colored cover pages. Utilization was assured by a subtle intermixing of elements which began with and which maintained organizational/political support. The Interim Report team planned, conducted, and reported its work with a stance which acknowledged that organizational and political support intermix; all evaluation activities were conducted within that constraint.

Recognizing the press on SEA politicos to justify their activities to legislators and recognizing legislators' anger with past evaluations, they chose a methodology which was timed for policymakers' needs and spoke the language of legislators (Dickey, 1980, has discussed the idea that evaluators and legislators are two communities who have different languages, timing needs, and dissemination modes). The case study approach allowed description of processes which relieved legislators from having to translate from technical language. The evaluation-as-disseminator mode allowed dissemination with explanation to many areas of education policymaking. It allowed for dissemination modes which acknowledged the power of "street-level bureaucrats" (Weatherly and Lipskey, 1977).
By providing policy implications to LEA administrators and teachers, and to SEA implementors, it acknowledged that the implementors of policy are as important in the policy process as the state legislator. It acknowledged the power of these implementors to transform policy and, in the process, become informal policymakers. It recognized that the evaluation should provide information to informal, as well as formal policymakers. In addition, they stay attuned to shifting needs and concerns of politicos. They maintain a policy-related focus in a policy-shifting environment; they also keep politicos informed so their agendas could anticipate the evaluation findings. Thus, this

The Intermixing of Crucial Elements for Facilitating Successful Evaluation Utilization

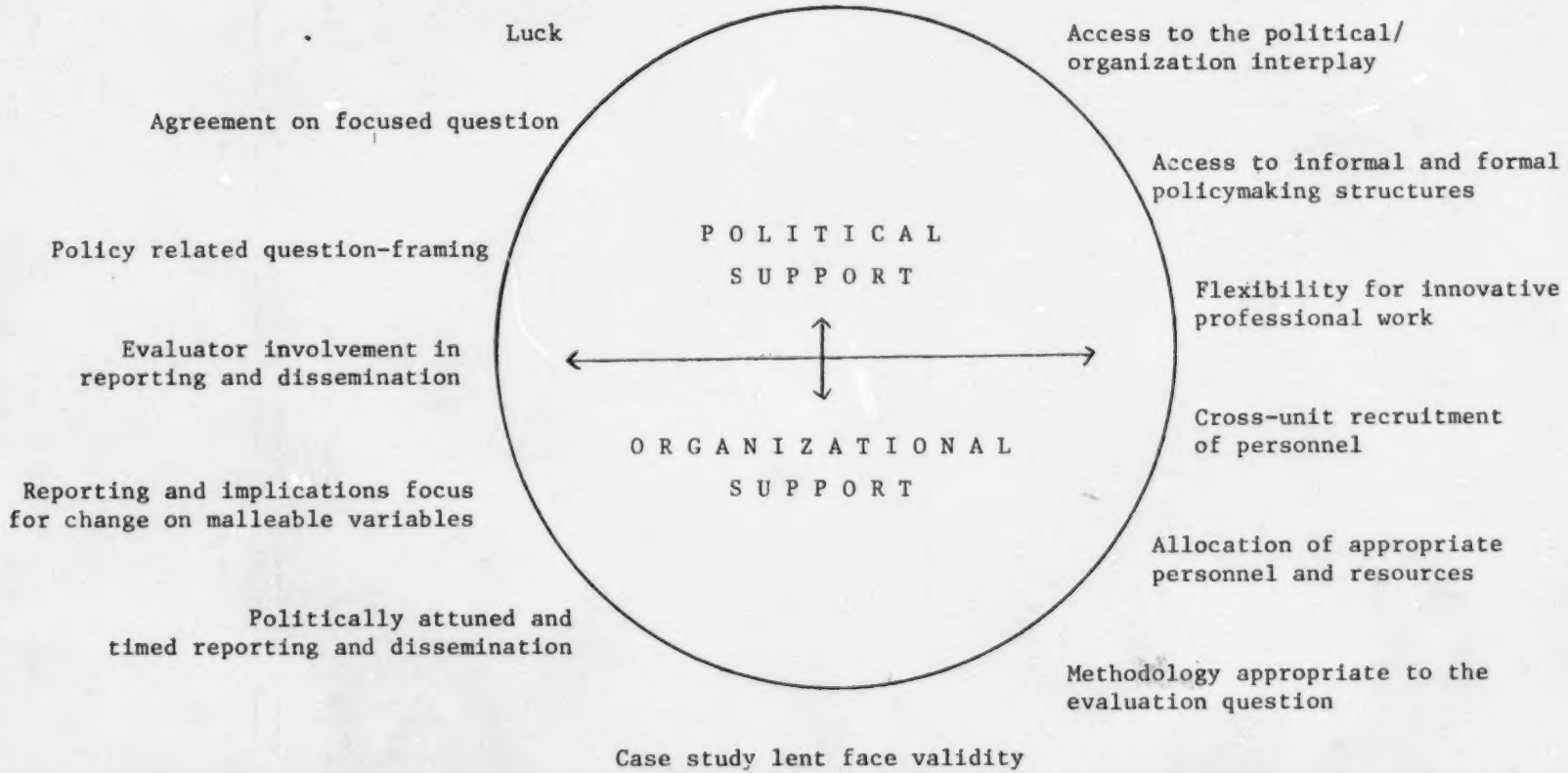
Luck
Agreement on focused question
Policy related question-framing
Evaluator involvement in reporting and dissemination
Reporting and implications focus for change on malleable variables
Politically attuned and timed reporting and dissemination
Case study lent face validity
Methodology appropriate to the evaluation question
Allocation of appropriate personnel and resources
Cross-unit recruitment of personnel
Flexibility for innovative professional work
Access to informal and formal policymaking structures
Access to the political/ organization interplay
POLITICAL SUPPORT
ORGANIZATIONAL SUPPORT

FIGURE THREE

evaluation was attuned to the education policy system in all its complexity.

The Interim Report was planned around a policy related, focused question which the various evaluation audiences agreed upon. This was good fortune, in part, but the evaluation team selected and defended a problem definition which focused on organizational and interorganizational processes and, from there, selected the appropriate case study methodology. They successfully trained themselves, explained and defended the methodology, anticipating the resistance to this innovative approach. They created a professional team spirit, a creative defensiveness and determination to do quality, significant work and ensure its utilization.

The evaluation team came together at a time of SEA organizational change--roles were altering, unit functions were shifting. The team's ability to recruit personnel from various SEA units and to manage themselves enabled them to create a salubrious team work environment. Situational leadership theory encourages such practices, where mature professionals work out their own structures for task division, time management, goal setting, recruitment and training (Hersey and Blanchard, 1977). This team earned this flexibility by staying attuned to organizational/political needs for information, simply through Daro's frequent informal reporting to SEA politicos. Additionally, the constant reporting prepared SEA politicos to create a constructive defense for any negative evaluation findings.

Cross-unit commitment to the utility and validity of the evaluation was a positive outcome of the cross-unit recruiting for the evaluation team. The usual intraorganizational controversies were alleviated and the evaluation incorporated all of their perspectives.

The report's style--mimeographed and stapled together--was less than slick. However, it was appropriately brief, it contained six pages of clear explanation of the methodology which answered the anticipated concerns for generalizability, validity, quality control and fears of legislators who were socialized to trust only statistics. Enough description was included to show the logic of conclusions and implications, but not so much as to make the report tedious. Finally, the report's policy implications were direct, specific, and aimed at all levels of the complex education policy system. They identified malleable variables in the system's processes and they provided suggestions for change. Their statements of "implications" were negotiated to avoid organizational resistance.

How did all of these key elements get interwoven to produce an evaluation which avoided political and organizational tensions and which answered needs of various audiences?--Luck, or at least taking advantage of a lucky setup. The time was an evaluation window; the variables were aligned. The SEA's constructive stance--the legislators' identifying a specific evaluation question, the SEA readying for new policy initiatives, SEA units realigning--created the evaluation window. Intraorganizational blockages loosened so that people with access, power, and expertise could work together. The SEA, legislators, LEAs and OPER were open to evaluation innovations.

Not every evaluator can arrange such evaluation windows. However, this evaluation analysis shows key elements which can be structured into evaluation management to facilitate evaluation utilization.

This case study of evaluation management confirms and explicates the analyses of previous reviewers of evaluation utilization mentioned earlier. However, this instance strongly emphasizes and illustrates the importance of

maintaining organizational/political support. It illustrates too that organizational/political support is a complex, interwoven system in education policy systems. Evaluation managers must respond to this.

References

Alkin, M., Daillak, R., & White, P. Using evaluations. Beverly Hills: Sage Publications, 1979.

Allison, G. Essence of decision: explaining the Cuban missile crisis. Boston: Little, Brown & Co., 1971.

Aufderheide, J. State policy making for the public schools of California. Columbus: Educational Governors Project, 1974.

Baker, E. Evaluation of the California Early Childhood Education Program, Volumes I-IV. Los Angeles: Center for the Study of Evaluation, UCLA, 1976.

Campeau, P. L. Rapid feedback evaluation of the Women's Educational Equity Act Program: summary. Palo Alto: American Institutes for Research in The Behavioral Sciences, 1981.

Cohen, D. K. Politics and research: evaluation of social action programs in education. Review of Educational Research, 1970, Vol. 40, No. 2, 213-238.

Cohen, L. Factors affecting the utilization of mental health evaluation research findings. Professional Psychology, 1977, Vol. 8, No. 4, 526-534.

Cook, T. D., & Reichardt, C. S. Qualitative and quantitative methods in research. Beverly Hills: Sage Publications, 1979.

Dickey, B. Utilization of evaluations of small-scale innovative educational projects. Educational Evaluation and Policy Analysis, 1980, Vol. 2, No. 6, 65-77.

Drew, E. Senator. New York: Simon & Schuster, 1979.

Edelman, M. The symbolic uses of politics. Urbana: University of Illinois Press, 1964.

Hershey, P., & Blanchard, K. H. Management of organizational behavior: utilizing human resources. Englewood Cliffs: Prentice-Hall, 1977.

Hill, P. Speech at UCLA, Managing the evaluation of Title I, Los Angeles, November, 1979

House, E. R. The conscience of education evaluation. Teachers College Record, 1972, Vol. 73, No. 3, 405-415.

Lindblom, C. The policy-making process. Englewood Cliffs: Prentice-Hall, Inc., 1968.

McLaughlin, M. W. Evaluation and reform. Cambridge: Ballinger, 1975.

March, J. G., & Simon, H. A. Organizations. New York: John-Wiley & Sons, Inc., 1958.

Marshall, C. Evaluation as a political tool: a case study of California Early Childhood Development Study. AERA Annual Conference, Los Angeles, 1981.

Mitchell, D. E. Social Science utilization in state legislative policymaking, Final report (NIE-G-76-0104). Washington, DC: National Institute of Education, 1979.

Parlett, M., & Hamilton, D. Evaluation as illumination: a new approach to the study of innovatory programs. From Gene Glass (ed.), Evaluation Studies Renew Annual, Vol. 1. Beverly Hills: Sage Publications, 1976.

Patton, M. Utilization-focused evaluation. Beverly Hills: Sage Publications, 1978.

Rosenthal, A. Utilization of evaluations in legislative settings. Paper presented at Evaluation Network Conference, Aspen, 1978.

Sroufe, G. E. Evaluation and politics, chapter X in J. Scribner, ed., The politics of education. Seventy-seventh yearbook of N.S.S.E., 1977.

State Department of Education, California. Interim report on the special study of selected ECE schools with increasing and decreasing reading scores. Sacramento, 1977.

State Department of Education, California. Consolidated application evaluation report. Sacramento, 1978.

State Department of Education, California. Report on the special studies of selected ECE schools with increasing and decreasing reading scores. Sacramento, 1980.

Weatherly, R., & Lipsky, M. Street-level bureaucrats and educational innovation. Harvard Educational Review, 1977, Vol. 47, 171-197.

Weiss, C. Evaluation research. Englewood Cliffs: Prentice-Hall, Inc., 1972.

_____ Where politics and evaluation research meet. Evaluation, 1973, Vol. 1, No. 3.

_____ Using social research in public policy-making. Lexington: Lexington Books, 1977.

_____ Knowledge creep and decision accretion. Knowledge: Creation, Diffusion, Utilization, 1980, Vol. 1, No. 3, 381-404.

Wilson, S. Explorations of the usefulness of case study evaluations. Evaluation Quarterly, 1979, Vol. 3, No. 3, 446-459.

Zeigler, E., & Valentine, J. Project Head Start. New York: The Free Press, 1979.

Footnotes

[1]There are many arguments to be made in favor of combining the case study approach with quantitative approaches, as argued in Cook and Reichart's volume, Quantative and Qualitative Methods in Evaluation Research (1979). The "rapid feedback evaluation" approach used to report to Congress on the Women's Educational Equity Act (Campeau, 1981) should be a useful approach to obtaining quick and relatively clean (as opposed to quick and dirty) evaluations of implementation.

[2]Quotes are identified by date only; they are derived from the research as described on page 4. Many in this paper are from members of The Interim Report team but the larger case study originally identified the degree and variety of utilization of this evaluation.

[3]Aufderheide's (1974) case study of California's education policy arena notes Riles' post-election staff changes--at top levels. In this research, 1979-1980 interviews indicate that, from the perspective of SEA insiders and others, the capabilities of top SEA policy personnel to creat policy outstrip the capacities of SEA staff to implement them.

[4]Others felt that Law had been constrained by the fact that he did not have staff for case studies and that he was a laissez-faire administrator who was unwilling to assertively promote a new approach to evaluation.

[5]Daro's original version had been twice as long.

[6]Daro and Denham skimmed through Chapter IV with me, checking off what was started, what was considered, and what was never done.

[7]SEA staff are discouraged from communicating with legislators and attending legislative hearings.

[8]The Interim Report on the Special Study of Selected ECE Schools with

Increasing and Decreasing Reading Scores (State Department of Education, 1977) was dissiminated as a xeroxed and stapled report. The pressure to publish the report waned since the findings were already being incorporated in policies within the SEA and in legislation. The final report was to include The Interim Report and the SRI Classroom Processes Study. Since Jane Stallings had already reported the SRI study in person to the legislature, and since The Interim Report findings had been published (an an interim report) and disseminated, the final publication was not crucial. Chapter II on statistics loomed as an embarassing appendage. In fact, an editor refused approval because it seemed so illogical. The Government Liaison person would not countenance publishing a report that differed from the original report which had gone to legislators; memories of the 1976 CSE Executive Summary fiasco still haunted (see p. 10). The final version, entitled entitled Report on the Special Studies of Selected ECE Schools with Increasing and Decreasing Reading Scores, was released in March, 1980. The controversial Chapter II had been deleted; Chapter IV, Department Action, did not appear in the final version. The Classroom Process Study, the new Chapter III, was introduced as a complementary research strategy, "a traditional correlational study of variables within classrooms" (SEA, 1980, p. 29).

[9]A high level LEA utilization could be assumed also, since the personal dissemination elicited high interest from LEAs, since the report included policy implications for LEAs, and since so many people requested copies.